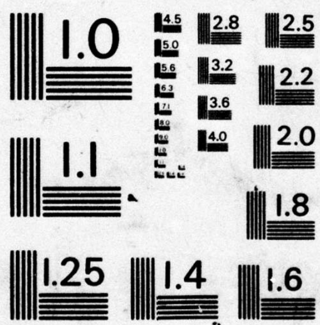
AD-A033 036 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 14/5
PIECEWISE APPROXIMATION OF PICTURES: FURTHER EXPERIMENTS.(U)
JUL 76 N AHUJA, L S DAVIS, D L MILGRAM F44620-72-C-0062
UNCLASSIFIED TR-462 AFOSR-TR-76-1135 NL

1 OF 1
AD
A033036

END
DATE
FILMED
1-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR - TR - 76 - 1135

COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

TR-462 July 1976
F44620-72C-0062

PIECEWISE APPROXIMATION OF PICTURES: FURTHER EXPERIMENTS

Narendra Ahuja
Larry S. Davis
David L. Milgram
Azriel Rosenfeld
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

In an earlier report, a new method of generating "skeleton" representations of noisy, grayscale pictures was described. This supplemental report investigates some variations on the basic method, and also shows how approximations to the picture can be constructed, given its "skeleton".

See 1423

The support of the Directorate of Mathematical and Information sciences, U.S. Air Force Office of Scientific Research, under Contract F44620-72C-0062, is gratefully acknowledged, as is the help of Mrs. Shelly Rowe in preparing this paper.

1. The SPAN

In [1], a general method of constructing "skeleton" representations of noisy, grayscale pictures was developed. This method assumed that a picture consists of a set of regions, each having approximately constant gray level (possibly noisy). The representation, referred to as the SPAN (for Spatial Piecewise Approximation by Neighborhoods), is constructed as follows:

1) At each point of the picture, we examine a set of neighborhoods of various sizes. For each neighborhood, a statistical test is applied to decide whether the neighborhood lies in a single region, or overlaps more than one region. Let N(x,y) be the largest neighborhood of (x,y) that lies in a single region.
2) N(x,y) is called maximal if it is not contained in any other (larger) N(u,v). The set of maximal N(x,y)'s is called the SPAN of the given picture.

The SPAN is a generalization, to grayscale pictures, of Blum's Medial Axis Transformation (MAT); see [1] for a more detailed comparison. It has several possible applications:

a) Construction of approximations to the picture
b) Smoothing the picture (replacing the gray level at each point (x,y) by the average gray level over N(x,y))
c) Detecting edges (at points where two maximal N(x,y)'s touch)

This report supplements [1] in several respects. It studies alternative statistical tests that can be used to choose N(x,y) at each point. It also develops methods of reconstructing approximations to a picture from its SPAN.

2. Tests for neighborhood acceptance

The statistical test used in [1] to determine whether a neighborhood lies in a single region was based on constructing a confidence interval around the neighborhood mean within which the true (region) mean should lie. If the neighborhood does lie within a single region, this interval should be small, whereas if it overlaps several regions, the interval should be large.

In this section we investigate some other methods of deciding whether or not a neighborhood lies in a single region. One approach is based on the assumption that the gray levels in each region are normally distributed; thus our decision criterion can take the form of a normality test. A second approach assumes that the gray level distribution in each region is unimodal, and takes the form of a multimodality test.

2.1 Normality Tests

A standard method of testing a distribution p(z) for normality is to divide the z-axis into intervals corresponding to equal areas under the normal curve (having the given mean and standard deviation). If the observed population is close to normal, it should be evenly distributed among these intervals. A chi-square test can be used to accept or reject this hypothesis at a given level of confidence.

The number of intervals used in this approach must be carefully chosen. If there are too few, the test may be too crude; but if there are too many, so that each one contains

only a few samples, the distribution of samples among the intervals may be uneven. We have chosen to make the number of intervals proportional to the sample standard deviation. It is also desirable to make this number a divisor of the sample size (here: the area of the neighborhood), to avoid restrictions on the degrees of freedom with which a point can go into any given interval. The number of intervals must be larger than 3, since the degrees of freedom of the chi-square test are reduced by 3 (because the mean, standard deviation, and number of classes are all known).

Let N be the number of intervals, and let n_i be the observed number of samples in the ith interval. If the expected number of samples in the ith interval for a $2k+1$ by $2k+1$ neighborhood is $(2k+1)^2/N$, then we have

$$\chi^2 = \frac{\sum_{i=1}^{N} [n_i-(2k+1)^2/N]^2}{(2k+1)^2/N}$$

Depending on the number of intervals used, we accept or reject the hypothesis that the neighborhood has a normal distribution according to whether the value of χ^2 is less or greater than the critical value at the desired confidence level.

In our experiments, the sample (= neighborhood) sizes used were $(2k+1)^2$ for $k = 1,\ldots,5$, and the confidence level used for rejecting normality was 95%. The number of intervals used was taken to be 1.8σ; this was rounded down to the nearest divisor of the neighborhood size. For $k=1$,

since the sample size was quite small, a confidence level of 80%, rather than 95%, was used. If the normality hypothesis was rejected even for k=1, the default decision was k=0, i.e., N(x,y) was taken to be {(x,y)} itself.

A supplementary test for normality was based on the skewness of the sample distribution. The skewness coefficient $\sqrt{B}$ for a sample (x_i's) with mean $\bar{x}$ is

$$\sqrt{B} = \frac{\sum(x_i-\bar{x})^3/(2k+1)^2}{[\sum(x_i-\bar{x})^2/(2k+1)^2]^{3/2}} = \frac{\sum(x_i-\bar{x})^3\cdot(2k+1)}{[\sum(x_i-\bar{x})^2]^{3/2}}$$

A neighborhood survives the skewness test if its $\sqrt{B}$ is below the critical value for the desired degree of confidence (we used 95%).

The skewness and χ^2 tests provide complementary information about possible non-normality of the sample distribution. For example, a neighborhood just on the border between two regions having the same variance would have a symmetrical distribution that would pass the skewness test, but that should be rejected by the χ^2 test. On the other hand, a neighboorhood that protrudes slightly across a region border might pass the χ^2 test, but would have an asymmetrical distribution that should be rejected by the skewness test. We required that both tests be passed in order to accept a neighborhood.

Figure 1 shows the original images (noise-free and noisy); the value of N(x,y) at each point (k=0, 1,...,5 represented by gray levels 10, 20,...,60); the values of the

maximal N(x,y)'s, with nonmaxima set to zero; the edges at which pairs of maximal N(x,y)'s meet; and the results of smoothing the images by replacing the gray level at each point (x,y) by the average gray level over the neighborhood N(x,y). These results are analogous to those in Figures 4, 5, 7, and 9 of [1]; the original images are the same as those in Figure 1 of [1]. It can be seen that

a) The skeletons (= sets of centers of maximal N(x,y)'s) are now much thinner, since a larger set of radii is being used ([1] allowed only k = 0, 1, 2)

b) Otherwise, the quality of the results is about the same as in [1].

2.2 Multimodality test

A much simpler approach to neighborhood acceptance is to test the distribution of gray levels in the neighborhood for multimodality. A neighborhood contained within a single region should have a unimodal distribution; thus if the distribution is found to be multimodal, we can assume that it does not lie in one region.

To test for multimodality, we first smoothed the neighborhood's gray level histogram by averaging over a 5-gray-level neighborhood of each histogram point. In the resulting smoothed histogram, changes in the sign of the slope that persisted for three or more successive points were counted. Each such change from positive to negative represents a peak. A histogram with two or more peaks was taken

to be multimodal.

Figure 2 is analogous to Figure 1, but using the multi-modality test. The skeletons and edges tend to be somewhat thinner and sharper than those in Figure 1, but there are more gaps. The smoothings of the noisy images are somewhat noisier. However, the overall performance is quite comparable to that using the more complex normality tests.

3. Image reconstruction from the SPAN

One of the useful properties of Blum's MAT is that it can be used to reconstruct the original image subset, since this is just the union of the maximal neighborhoods in the MAT. More important, if neighborhoods of small sizes are omitted from the MAT, the reconstruction yields a simplified version of the original subset, with small parts deleted but major parts (approximately) preserved. This section addresses the corresponding problem of reconstructing an approximation to the original image from the SPAN.

The natural building blocks for such a reconstruction are the maximal neighborhoods N(x,y), each filled in with a constant gray level equal to its average gray level on the original image. However, it is not clear how these N(x,y)'s should be combined when they overlap (and have different gray levels). We have chosen, somewhat arbitrarily, to use the following rules of combination:

1) When maximal N(x,y)'s having the same radius overlap, we use the maximum of their gray levels. [Alternatives would be to use the minimum, or the average; all of these alternatives will be illustrated below.]
2) When maximal N(x,y)'s having different radii overlap, we use the gray level belonging to the one with the smaller radius. [Rationale: The ones with small radii are needed to provide fine image detail; they could not do this if they were overidden by the large ones. Indeed, the large ones will pass the acceptance tests even when they slightly overlap

neighboring regions, and this could cause the mean gray level of one region to be given to points of an adjacent region.]

Figure 3 shows steps in the reconstruction of the images of Figures 1-2 (both noisy and non-noisy) using these rules of combination. We first display the N(x,y)'s having each individual radius (k = 5, 4, 3, 2, 1, 0), and then display the cumulative effects of combining the N(x,y)'s having different radii, beginning with the largest radius (5+4; 5+4+3; 5+4+3+2; 5+4+3+2+1; 5+4+3+2+1+0). It is seen that

a) The reconstruction of the non-noisy images resemble the original images closely.

b) The reconstructions of the noisy images resemble the originals crudely, but have a mottled appearance. This is primarily caused by the 3-by-3 N(x,y)'s (k=1), for which an over-tolerant 80% confidence level was used.

c) The smallest size (k=0) contributes almost nothing to the reconstruction.

The N(x,y)'s obtained using the normality tests were used in this Figure.

The results are similar if we use N(x,y)'s obtained from the multimodality test, or if we use the average or min, rather than the max, in combining the N(x,y) gray levels. Final reconstructions (5+4+3+2+1+0) for all of these cases are shown in Figure 4. We see that

a) The results using max, min, and average are all quite similar.

b) The results based on the multimodality test are sharper, but more blocky, than those based on the normality tests. Apparently this test accepted too many 5-by-5 neighborhoods and too few 3-by-3's.

A more quantitative comparison of these results can be made with reference to Figures 5 and 6. Figure 5 tabulates the number of maximal N(x,y)'s of each radius, for each of the ten input pictures, based on both the normality and multimodality tests. It is seen that the multimodality test yields substantially fewer SPAN points than the normality tests.

Figure 6 tabulates the mean absolute reconstruction errors for each of the ten input pictures, using the two types of N(x,y)'s (based on normality and multimodality) and the three combination rules (max, min, average). It is seen that

a) These errors only amount to a few gray levels (on a gray scale of 0-63).

b) The errors are generally smaller for the multimodality test than for the normality test, in spite of the latter's involving more SPAN points.

c) For a given test, the average, min, and max reconstruction schemes do about equally well. (The average does noticeably better for the normality test in the case of the chromosome image, perhaps because its edges are so blurred.)

4. <u>Compression based on the SPAN</u>

If the pictures reconstructed from the SPAN are acceptable approximations to the original pictures, the SPAN can be used as the basis for a picture compression scheme. In fact, suppose that we encode the picture points as follows:

0 - nonSPAN point

$1\alpha\beta$ - SPAN point having radius α and mean gray level β in its maximal neighborhood.

Here α is a 3-bit number (the possible radii are 0, 1,...,5), and β is a 6-bit number (assuming that we round the average gray levels to the nearest integer in the original range 0,...,63). Thus the $1\alpha\beta$ codes have ten bits each (1+3+6). If the fraction of SPAN points in the picture is p, and the picture has N points, then the number of bits required to encode the picture in this way is

$$10pN + (1-p)N = (1+9p)N$$

On the other hand, if we simply encoded each picture point by its gray level (a 6-bit number), the picture would require 6N bits. Using the SPAN encoding is thus more economical provided $1+9p < 6$, i.e., provided $p < 5/9$.

As an example, consider the first of our ten pictures, for which the SPAN constructed using the normality test has 337 (out of N = 1024) points, while the SPAN based on the multimodality test has only 231 points. Here encoding each picture point by its gray level would require $6 \times 1024 = 6144$ bits. When the SPAN is used, however, the numbers of bits needed is

$$10 \times 337 + (1024-337) = 3370 + 687 = 4057$$

and $10 \times 231 + (1024-231) = 2310 + 783 = 3103$

for the normality and multimodality approaches, respectively; the latter represents a compression of nearly 2:1. The results are similarfor the remaining nine pictures, both noisy and non-noisy.

[If there were very few SPAN points, an alternative method of encoding would be to simply list the radii, gray levels, and coordinates of these points. In our cases, where the pictures are 32 by 32 (or smaller), the coordinates would require 10 bits (5 each for row number and column number), which together with the 9 bits for gray level and radius comes to 19 bits per SPAN point. If there were (say) only 100 SPAN points, this would require 1900 bits total, for a compression of over 3:1; but for 200 or more SPAN points it requires 3800 or more bits, which is more than were needed using the method described above.]

5. The SPAN as a segmentation aid

It is well known that a picture can often be segmented by detecting valleys in its gray level histogram, and choosing thresholds at the bottoms of these valleys. The SPAN approach can be used to make histogram valleys easier to detect. This is illustrated in Figure 7, which shows gray level histograms for the six original non-artificial pictures; the smoothed pictures (see Figure 1) using N(x,y)'s determined by both the normality and multimodality tests; and the reconstructed pictures (see Figure 4) using SPAN points chosen by both tests, and the max rule for combining gray levels.

It is seen that the smoothed pictures have sharper-peaked histograms than the original pictures, while the histograms of the reconstructed pictures are sharper still. Thus smoothing using the N(x,y) neighborhoods, and reconstruction using the maximal N(x,y)'s, tends to yield histograms that should be easier to segment.

Reference

1. L. S. Davis, A. Rosenfeld, and N. Ahuja, Piecewise approximation of pictures using maximal neighborhoods, Technical Report 455, Computer Science Center, University of Maryland, College Park, Md., May 1976.

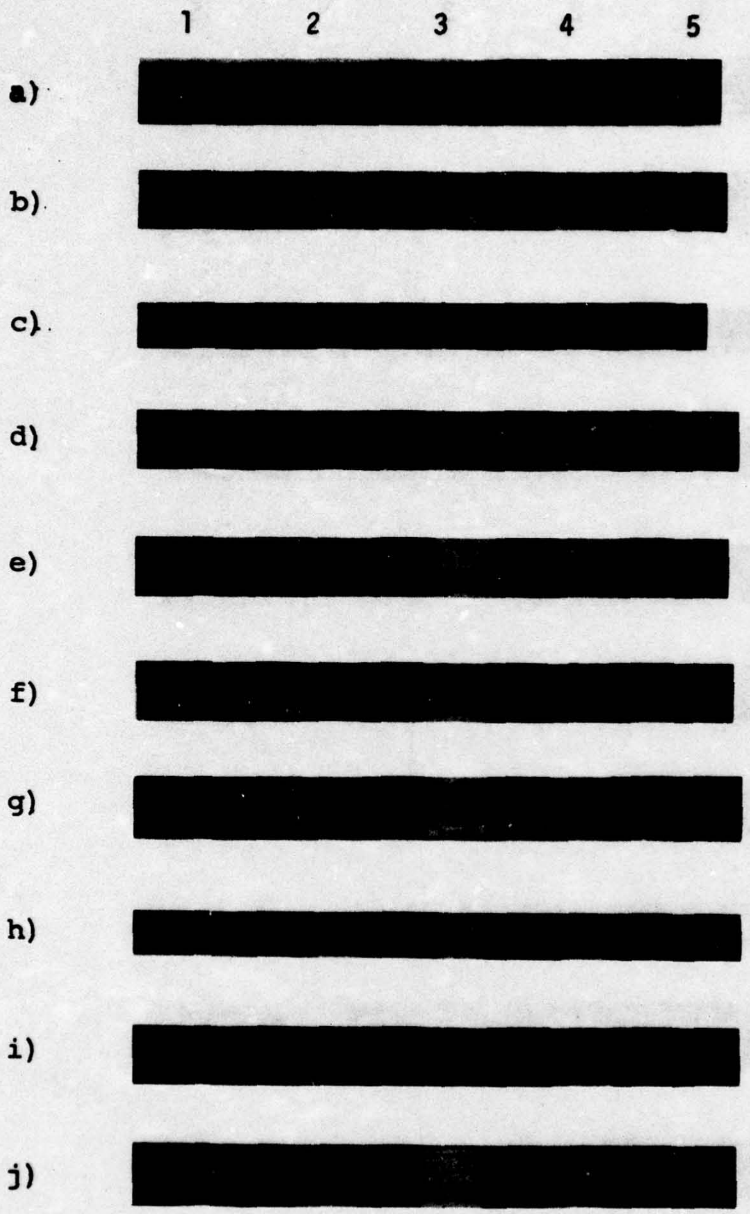


Figure 1. Results of neighborhood selection using normality tests:

(1) Original image; (2) N(x,y) radii; (3) maximal N(x,y) radii; (4) edges; (5) smoothed images. Images (a-e) are noise-free, (f-j) are noisy.

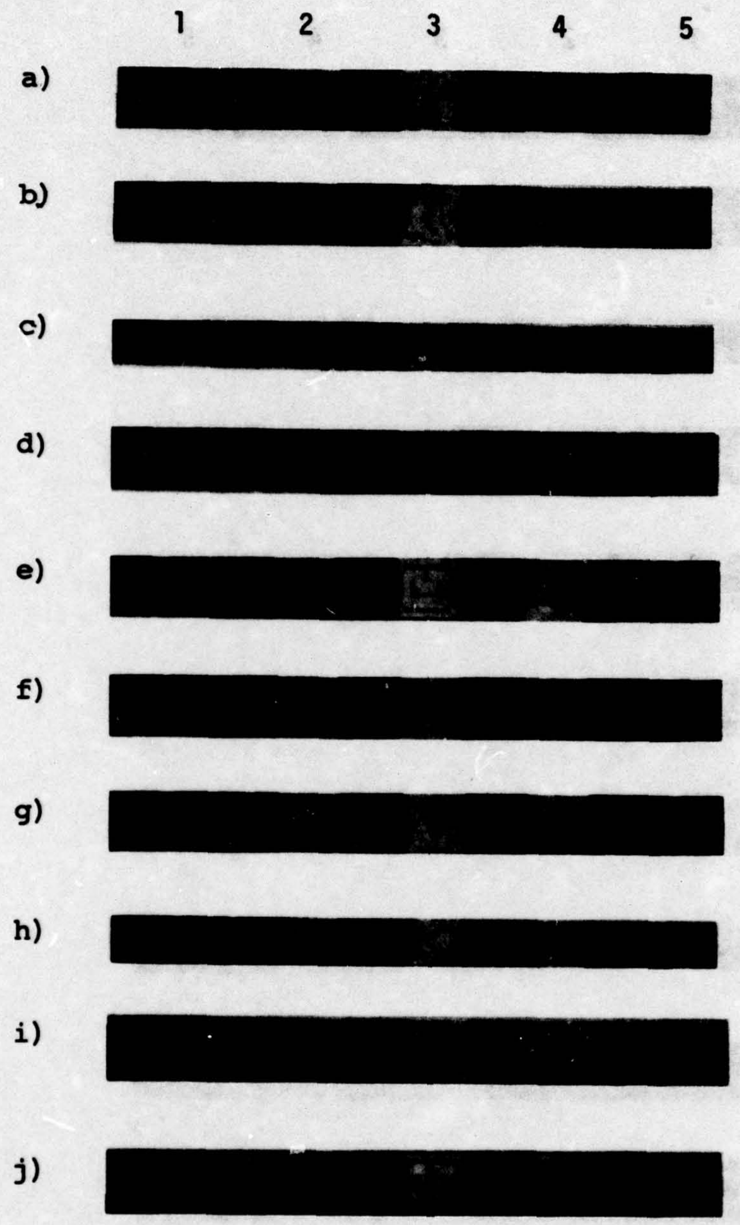


Figure 2. Analogous to Figure 1, but using multimodality test.

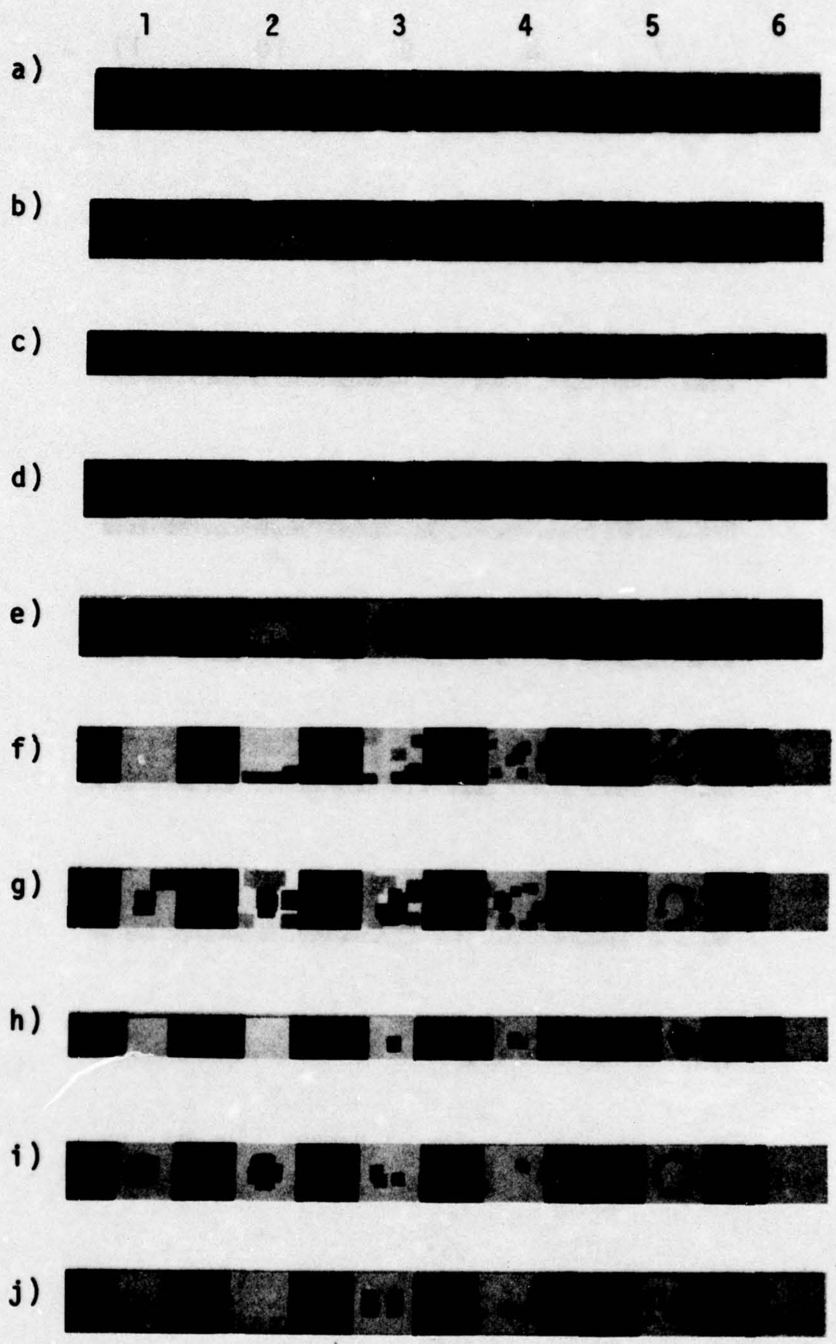


Figure 3. Steps in the reconstruction of the images (a-j) used in Figures 1-2 from their SPANS. (1-6) N(x,y)'s having k = 5, 4, 3, 2, 1, 0; (7-11) Cumulative displays for 5+4, 5+4+3, 5+4+3+2, 5+4+3+2+1, and 5+4+3+2+1+0

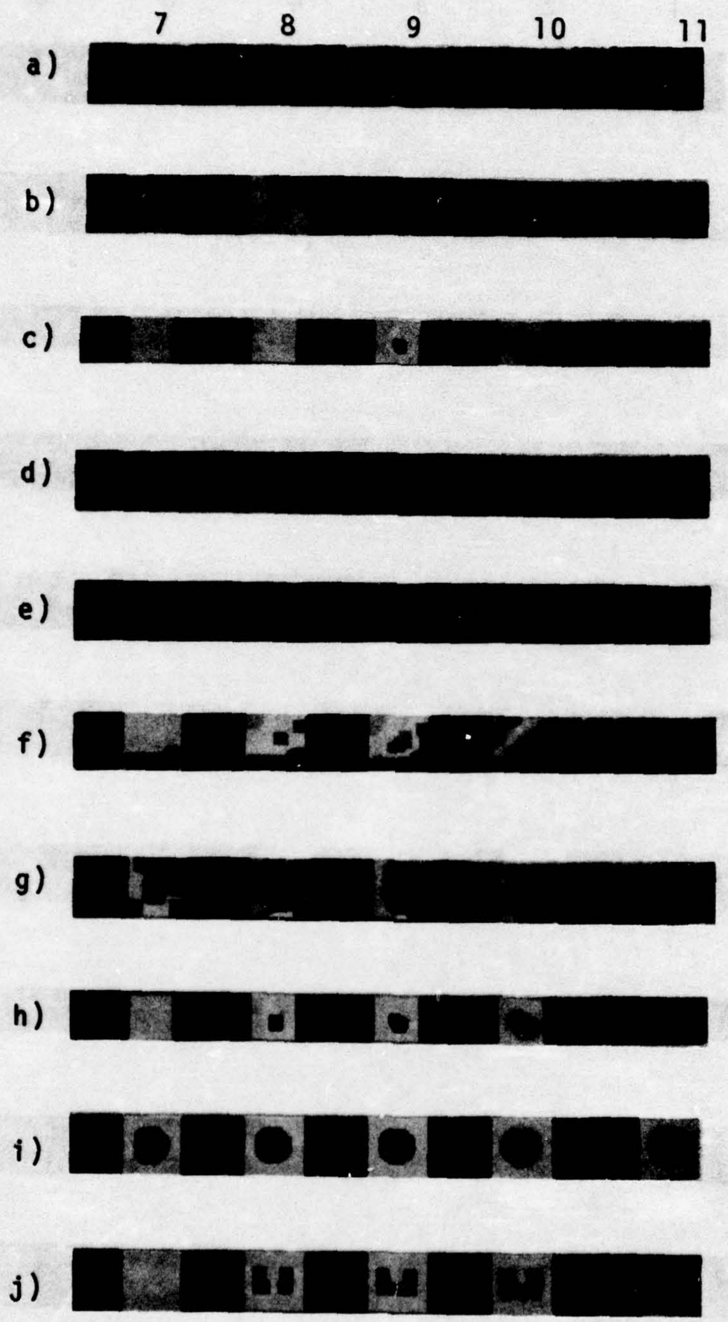
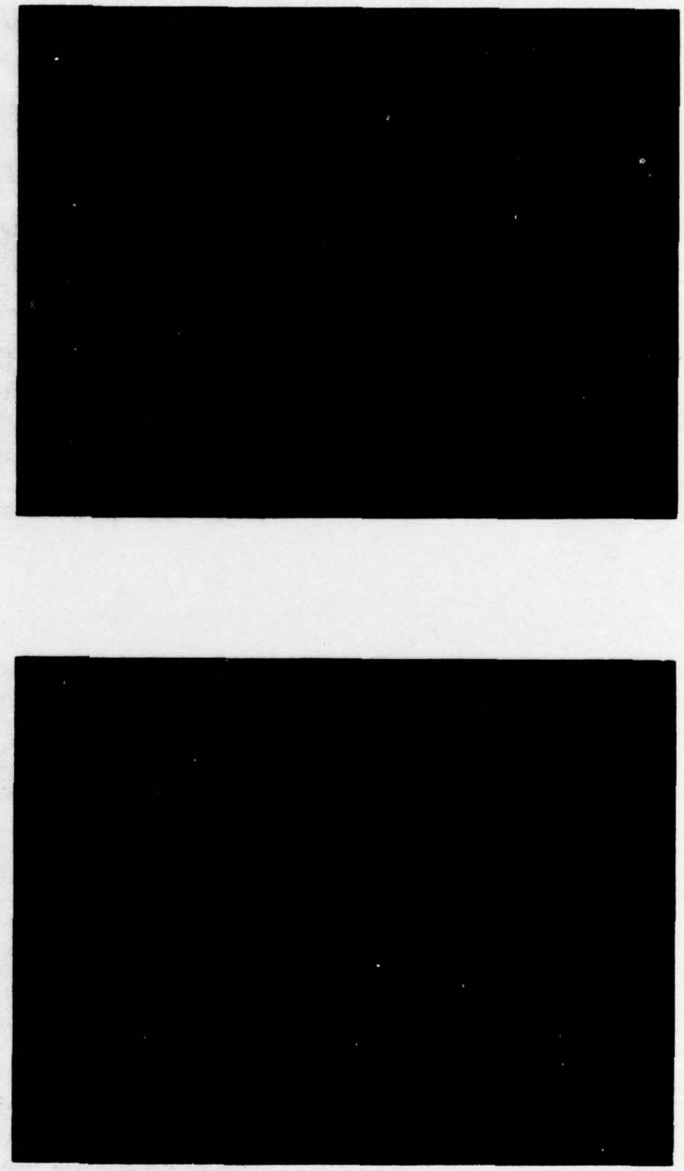


Figure 3 (Continued)

(1) max

(A)

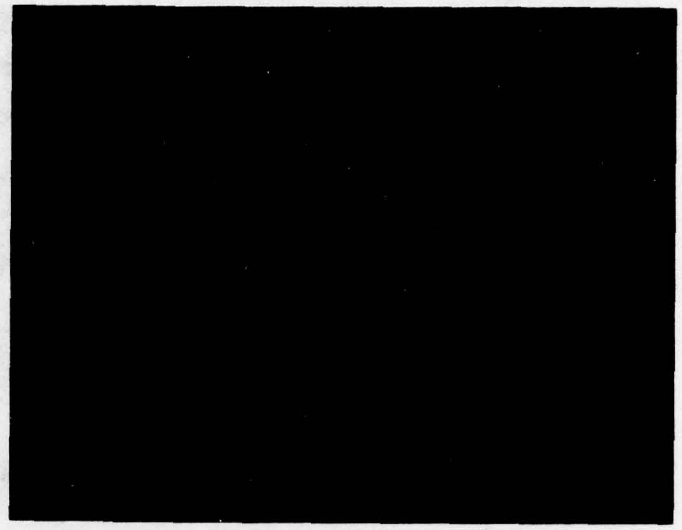


(B)

Figure 4. Reconstructions (5+4+3+2+1+0; see Figure 3) based on N(x,y)'s obtained from (A) normality and (B) multimodality tests, using (1) max, (2) min, and (3) average for combining gray levels.

(2) min

(A)



(B)

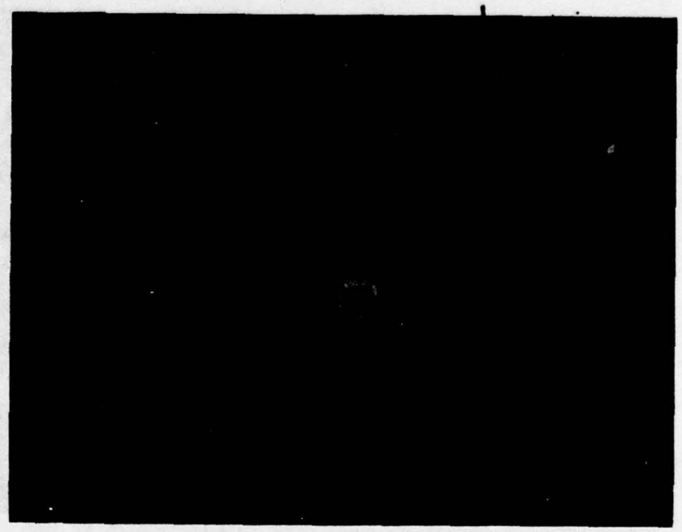
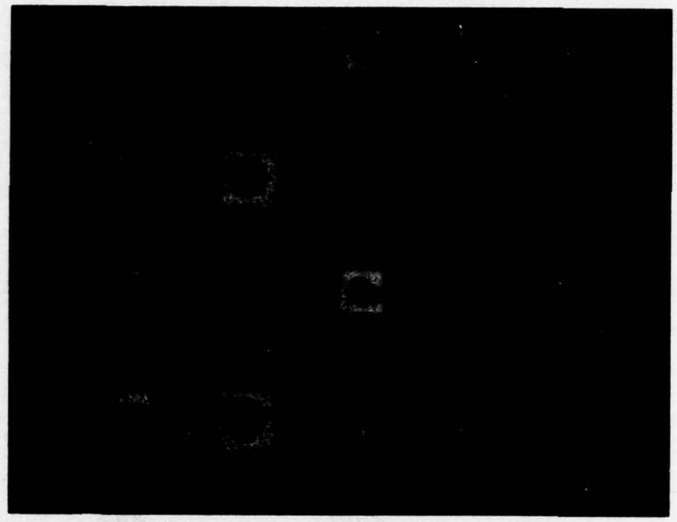


Figure 4 (Continued)

(3) average

(A)



(B)

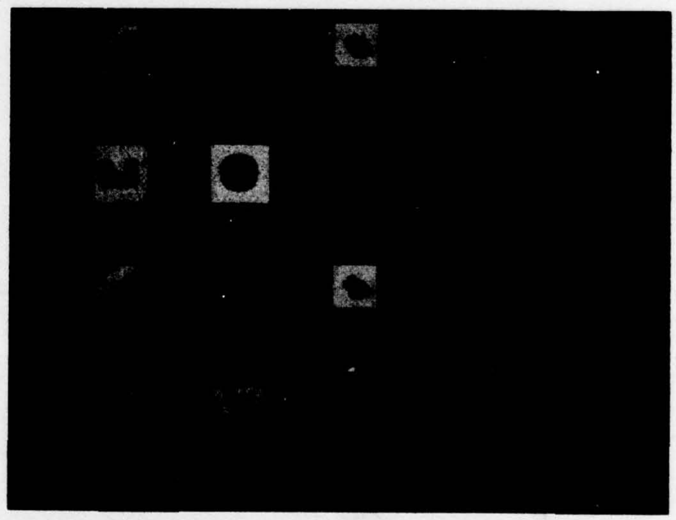


Figure 4 (Continued)

Picture	Radius	Number of points: Normality	Number of points: Multimodality	Number of SPAN points: Normality	Number of SPAN points: Multimodality
a	5	-	1	-	1
	4	2	26	2	25
	3	14	54	11	13
	2	85	139	62	55
	1	452	290	236	38
	0	471	514	26	99
	Total	1024	1024	337	231
b	5	-	18	-	18
	4	22	84	22	43
	3	49	150	16	33
	2	150	210	92	19
	1	537	256	250	32
	0	266	306	9	-
	Total	1024	1024	389	145
c	5	-	-	-	-
	4	-	9	-	9
	3	8	43	8	19
	2	67	131	45	53
	1	368	171	230	14
	0	133	222	-	23
	Total	576	576	283	118
d	5	-	-	-	-
	4	48	48	48	48
	3	144	144	36	36
	2	256	256	54	54
	1	284	284	-	-
	0	292	292	-	-
	Total	1024	1024	138	138

Figure 5. Number of SPAN points (maximal N(x,y)'s) of each radius for each of the input pictures (a-j), based on the normality and multimodality tests.

Picture	Radius	Number of points: Normality	Number of points: Multimodality	Number of SPAN points: Normality	Number of SPAN points: Multimodality
e	5	37	37	37	37
	4	41	41	9	9
	3	128	128	52	52
	2	254	254	62	62
	1	267	276	-	-
	0	288	288	4	4
	Total	1024	1024	164	164
f	5	-	1	-	1
	4	10	15	10	15
	3	51	34	27	12
	2	145	78	60	51
	1	542	163	281	31
	0	276	733	2	160
	Total	1024	1024	380	270
g	5	7	24	7	24
	4	41	40	23	25
	3	164	53	76	16
	2	213	57	34	12
	1	393	122	71	25
	0	206	728	-	56
	Total	1024	1024	211	158
h	5	-	-	-	-
	4	-	3	-	3
	3	21	24	21	14
	2	100	63	54	24
	1	281	89	139	27
	0	174	397	2	63
	Total	576	576	216	131

Figure 5 (Continued)

Picture	Radius	Number of points: Normality	Number of points: Multimodality	Number of SPAN points: Normality	Number of SPAN points: Multimodality
i	5	-	-	-	-
	4	31	48	31	48
	3	102	142	41	35
	2	204	238	52	47
	1	365	223	85	-
	0	322	373	-	-
	Total	1024	1024	209	130

Picture	Radius	Number of points: Normality	Number of points: Multimodality	Number of SPAN points: Normality	Number of SPAN points: Multimodality
j	5	17	37	17	37
	4	38	41	17	9
	3	95	126	35	52
	2	217	229	67	57
	1	338	215	84	-
	0	319	376	-	5
	Total	1024	1024	220	160

Figure 5 (Continued)

Picture	Using normality tests Max	Min	Average	Using multimodality test Max	Min	Average
a	3.59	3.79	3.27	2.37	2.53	2.40
b	2.29	2.62	2.16	2.72	2.92	2.70
c	3.64	3.60	2.16	2.40	2.44	2.34
d	0	0	0	0	0	0
e	0	0	0	0	0	0
f	4.95	5.58	4.90	2.96	3.16	3.10
g	2.98	3.48	3.09	2.62	2.79	2.67
h	4.06	3.85	2.88	2.38	2.49	2.40
i	3.58	3.76	3.79	0.80	1.62	1.27
j	2.89	3.13	3.10	0.99	1.30	1.25

Figure 6. Mean absolute reconstruction errors corresponding to the cases in Figure 4. Errors are measured in gray levels on a scale of 0 to 63.

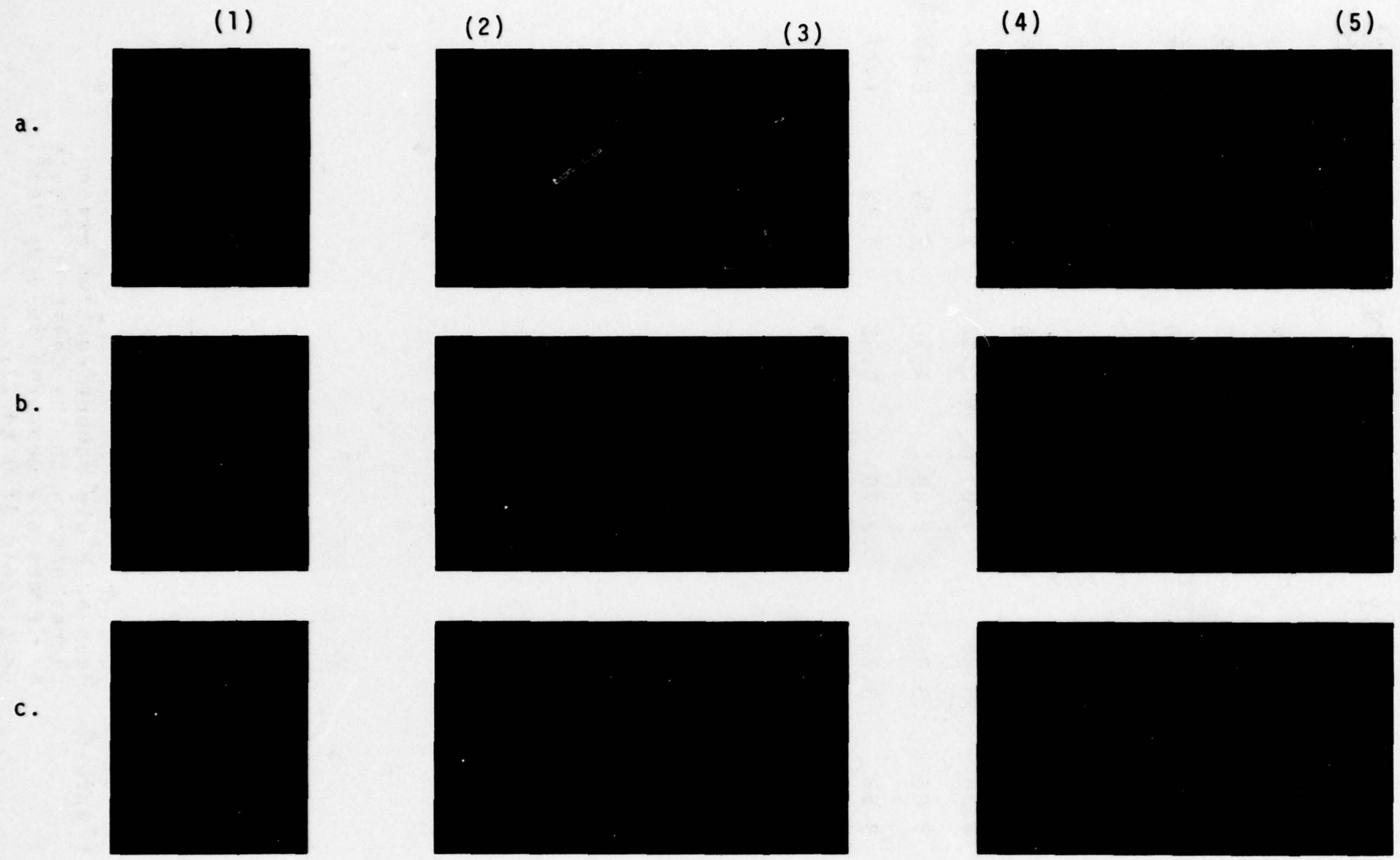


Figure 7. Histograms: (1) original; (2) smoothed, using normality test; (3) reconstructed, using normality test and max rule; (4) smoothed, using multimodality test; (5) reconstructed, using multimodality test and max rule.

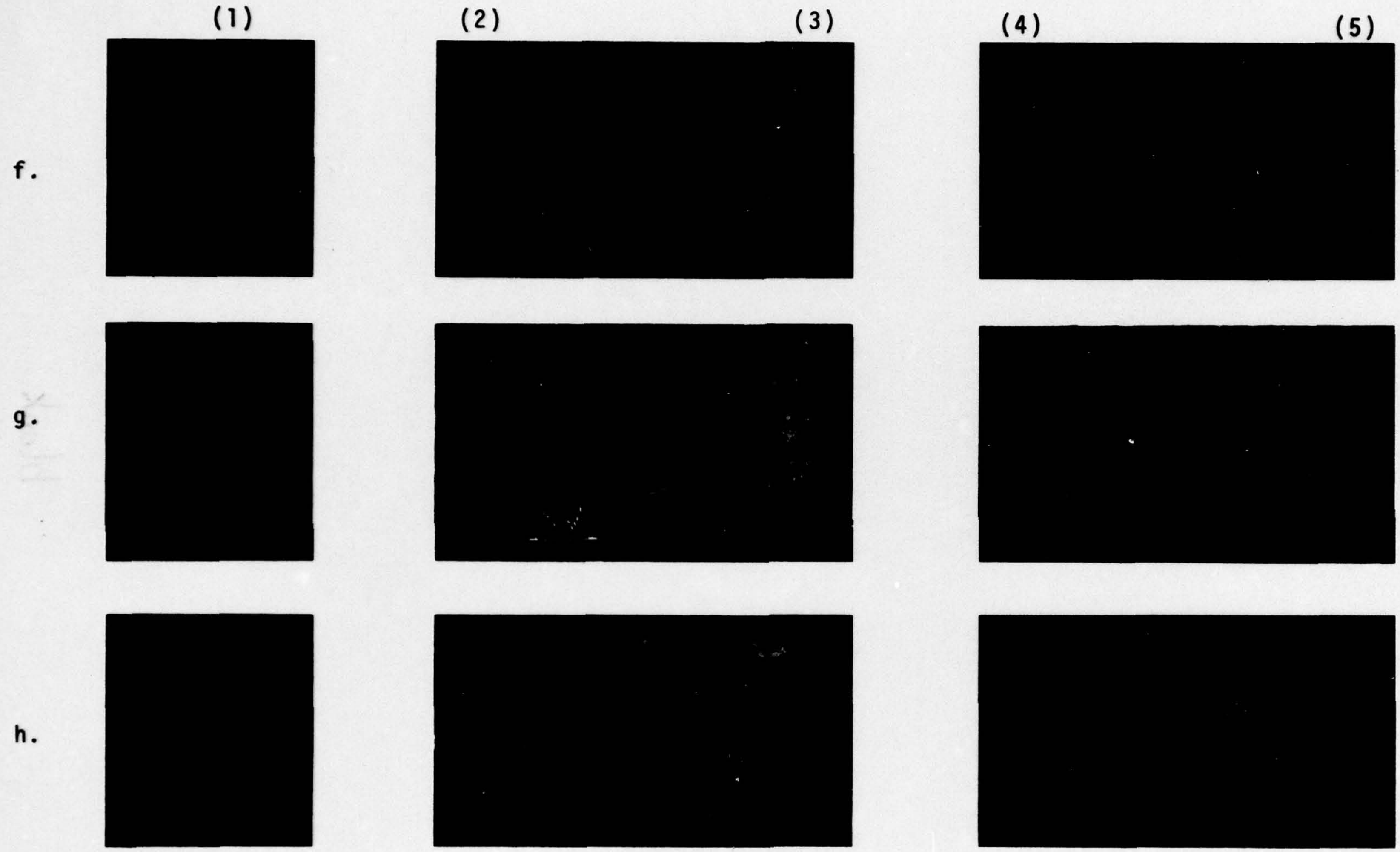


Figure 7 (Continued)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (19) (18) AFOSR - TR - 76 - 1135
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER (9) Technical rept.,
4. TITLE (and Subtitle) (6) PIECEWISE APPROXIMATION OF PICTURES: FURTHER EXPERIMENTS.
5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER (14) TR-462
7. AUTHOR(s) (10) Narendra Ahuja, Larry S. Davis, David L. Milgram [redacted] Azriel Rosenfeld
8. CONTRACT OR GRANT NUMBER(s) (15) F44620-72C-0062
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Computer Science Ctr. Univ. of Maryland College Park, MD 20742
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 9769-02
11. CONTROLLING OFFICE NAME AND ADDRESS: Math.& Info. Sciences, AFOSR/NM Bolling AFB Wash., DC 20332
12. REPORT DATE: July [redacted]76 (11) Jul 76
13. NUMBER OF PAGES: 28 (12) 29p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (16) 9769 (17) 02
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image processing
Pattern recognition
Medial axis transformations
Skeletons
Smoothing
Edge detection

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In an earlier report, a new method of generating "skeleton" representations of noisy, grayscale pictures was described. This supplemental report investigates some variations on the basic method, and also shows how approximations to the picture can be constructed, given its "skeleton".

403018